



JOSEPH BOWEN 

‘BUT YOU COULD HAVE HURT ME!’: RISK AND HARM

(Accepted 3 October 2021)

ABSTRACT. This paper answers two questions. First, on the assumption that risk of harm is of moral significance, does risk’s moral significance lay in its being harmful? Second, is risk of harm itself harmful? I argue that either risk is not harmful or that risk is harmful only in a small range of cases. If risk is not harmful, and yet risk is of moral significance, risk’s moral significance cannot lie in its being harmful. And if risk is harmful only in a small range of the cases in which risk is of moral significance, risk’s moral significance cannot lie in its being harmful.

I. INTRODUCTION

Let us begin with a case.

Roulette. Target is asleep. Her housemate, Shooter, comes into her room and plays Russian roulette with her. She pulls the trigger with the gun pointed at Target but, luckily, no bullet is fired. Shooter, content with having played a round of roulette, never plays roulette again.

Let us call this a case of *pure risk imposition*.¹ If, the following morning, Shooter were to explain to Target what happened, Target might reasonably become distressed. She might fear that Shooter will, again, break into her room and play Russian roulette with her, despite Shooter’s assurances to the contrary. She might buy a lock for her door, spend time at friends’ houses, or even move. In all of these cases, Target might be harmed by Shooter. But these harms (the distress, fear, and disruption) are *downstream* from the risk imposition and, as *Roulette* makes salient, contingent. While of moral significance, they are not the subject of our discussion.

¹ Judith Jarvis Thomson, ‘Imposing Risks’, in *Rights, Restitution, and Risk: Essays in Moral Theory*, ed. William Parent (Cambridge, MA: Harvard University Press, 1986), 176.

Everyone should agree that what Shooter does is of moral significance. I am going to proceed on this assumption.² My interest in this paper lies in *why* it is of moral significance. One popular answer is that her action is morally significant because it is itself harmful.³ To assess this answer, we need to determine whether risk of harm is itself harmful. This is all to say, we need to ask two questions. First, assuming risk of harm is of moral significance, does risk's moral significance lay in its being harmful? Second, is risk of harm itself harmful? (For brevity, is risk harmful?)

I argue that risk is not harmful. If risk is not harmful, and yet risk is of moral significance, risk's moral significance cannot lie in its being harmful. For reasons that become clear below, some people might not follow me all the way to thinking that risk is *never* harmful. I argue that, even if this weaker view is correct, risk is nonetheless harmful only in a small range of cases. If risk is harmful only in a subset of the much larger set of cases in which risk is of moral significance, risk's moral significance cannot lie in its being harmful.

It would be useful both if risk were harmful and if risk's moral significance lay in its being harmful. This would be useful because it would make our moral, legal, and political theorising easier than if risk were not harmful. This is because harm already plays a significant role in our theorising. Let me offer three examples.⁴

² I am going to leave the sense in which risk is of moral significance intuitive. My hope is that everyone should be onboard with the view that what Shooter does matters morally. For those unsatisfied, below in the introduction I list three ways in which people think harm is of moral significance: it matters for criminalisation, the stringency of rights, and the proportionality constraint on defensive harming. On the face of things, I think risk of harm is relevant to these three domains, just as harm is relevant. As I explain below in the text, if risk is itself harmful, we have a ready explanation for why this is.

³ Claire Finkelstein, 'Is Risk a Harm?', *University of Pennsylvania Law Review* 151, no. 3 (2003): 963–1001; Matthew Adler, 'Risk, Death, and Harm: The Normative Foundations of Risk Regulation', *Minnesota Law Review* 87, no.5(2003): 1293–1445; John Oberdiek, 'Towards a Right Against Risking', *Law and Philosophy* 28, no. 4 (2009): 367–392; John Oberdiek, 'The Moral Significance of Risking', *Legal Theory* 18, no. 3 (2012): 339–356; John Oberdiek, *Imposing Risk: A Normative Framework*, Oxford Legal Philosophy (Oxford: Oxford University Press, 2017); Seth Lazar, 'Risky Killing and the Ethics of War', *Ethics* 126, no. 1 (2015): 91–117; Seth Lazar, 'Risky Killing: How Risks Worsen Violations of Objective Rights', *Journal of Moral Philosophy* 16, no.1 (2017): 1–26; Adriana Placani, 'When the Risk of Harm Harms', *Law and Philosophy* 36, no. 1 (2017): 77–100. With the exception of, perhaps, Lazar, all seem to imply that the moral significance of risk lies in its being harmful.

⁴ Cf., e.g., Bradley, who argues that harm should not play such an important role in our theorising: Ben Bradley, 'Doing Away with Harm', *Philosophy and Phenomenological Research* 85, no. 2 (2012): 390–412. Note, I am not saying that risk's being harmful is *essential* to dealing with these three examples. I am saying only that it would make our theorising easier if risk were harmful. For that reason, I do not discuss ways of dealing with these three examples that do not appeal to the claim that risk is harmful.

First, many people think that harm plays a role in the legitimate exertion of power over others. For example, Joel Feinberg thinks that a sufficient condition for a consideration's counting in favour of penal legislation is that it prevents harm to others.⁵ Some read the harm principle differently. John Stuart Mill thinks that a necessary condition on restricting the liberty of a person is that what is restricted would have been harmful to others.⁶ If risk is harmful, we can easily account for why we can exert power over Shooter to stop her from playing Russian roulette with Target.

Second, in *Roulette*, intuitively we want to say that Target has a stringent – that is to say, very important – right against Shooter's subjecting her to risk.⁷ However, many theories of rights say that, other things being equal, the stringency of a right corresponds to the harm that would befall its holder, were that right not to be respected.⁸ For example, my right that you not subject me to extreme torture is much more stringent than my right that you not slap me. A plausible explanation of this is that it would be much worse for me were you to torture me than were you to slap me, and the stringency of my rights reflects this. What, then, is the stringency of Target's right that Shooter not play roulette with her?

Third, most theories of defensive harm say that some act of defensive harm is permissible only if the harm caused is proportionate to the harm averted. But if risk is not itself harmful, there is no harm to avert in cases of pure risk imposition. Yet, we might nevertheless want to say that Shooter is liable to be harmed in defence of Target. We might also want to say that Shooter is not permitted to defend herself, if someone attempts to defend Target

⁵ Joel Feinberg, *Moral Limits of the Criminal Law. Volume 1, Harm to Others* (Oxford: Oxford University Press, 1984).

⁶ John Stuart Mill, 'On Liberty', in *The Collected Works of John Stuart Mill*, vol. XVIII (Toronto: University of Toronto Press, 1977), 213–310.

⁷ David McCarthy, 'Rights, Explanation, and Risks', *Ethics* 107, no. 2 (1997): 205–225; Oberdiek, *Imposing Risk*, *supra* note 3; Michael J. Zimmerman, *Ignorance and Moral Obligation* (Oxford: Oxford University Press, 2014); Stephen Perry, 'Torts, Rights, and Risk', in *Philosophical Foundations of the Law of Torts*, ed. John Oberdiek (Oxford: Oxford University Press, 2014). While Thomson thinks we do not have rights against risk of harm, she sees the appeal: Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, MA: Harvard University Press, 1990), 242–248.

⁸ While those who endorse the Interest Theory of rights will be very amiable to this idea, many who are sceptical of the Interest Theory also endorse it, e.g., Thomson, *The Realm of Rights*, *supra* note 7 at 149–175; F. M. Kamm, *Intricate Ethics* (Oxford: Oxford University Press, 2007), 249–275. The "other things being equal" holds fixed, among other things, the way in which the person is harmed, and so is compatible with the view that rights against being harmed are more stringent than rights to be rescued from equivalent harms when all else is equal.

against Shooter. But if Shooter is not liable to be harmed, she has a right against being harmed, which, absent further argument, would be violated if she is killed in defence of Target. And usually, agents are permitted to defend themselves when their rights are violated (subject to that defensive harm satisfying the other conditions of permissible self-defence).

If risk is not harmful, there are plenty of things that can be said in reply to these three sketches of harm's importance. My point has been only to demonstrate that our theorising will be easier if risk is harmful than if risk is not harmful.⁹

Let me outline the structure of this paper. We see that there are at least two explanations for the idea that risk is harmful. First (section II), often people desire not to be the subject of risk. Perhaps people are harmed through being at risk because their desires are frustrated. Second (section III), by exposing others to risk, one might frustrate others' autonomy. Perhaps people are harmed through having their autonomy frustrated. I argue that both views are found wanting. In section IV, I briefly object to some additional explanations for the view that risk is harmful before, in section V, arguing against the view that risk is harmful, however its harmfulness is explained. I conclude by seeing how we might move forward in determining the moral significance of risk.

A preliminary. Let us proceed on the following standard *Comparative Counterfactual Account of Harm*: an event, *e*, harms a subject, *S*, iff *S* is worse off (all things considered) than she would have been had *e* not occurred. And let us again follow a standard line and say *S* is worse off than she would have been iff her wellbeing is lower than it would have been.¹⁰ We begin by being agnostic on what wellbeing consists in. Three things to note about the Counterfactual Account. First, it allows events caused by non-agents to be harmful. A tree's

⁹ Determining whether risk is harmful is also of significance because the law needs an answer to this question. The paradigm case is *Hotson v East Berkshire Area Health Authority* [1987] AC 750 (HL). See: Stephen Perry, 'Risk, Harm and Responsibility', in *Philosophical Foundations of Tort Law*, ed. David G. Owen (Oxford: Clarendon Press, 1995); Perry, 'Torts, Rights, and Risk', *supra* note 7 at 45–48; Finkelstein, 'Is Risk a Harm?', *supra* note 3 at 975–990.

¹⁰ We can separate questions concerning the *measure* of harm (whether we should be counterfactual comparativists, temporal comparativists, noncomparativists, and so on) from questions concerning the *currency* of harm (whether we should think harm concerns bad things happening to our wellbeing, our desires, and so on, where 'bad things' should be read loosely enough to be agnostic on the currency of harm). On the first question, see, among others, Bradley, 'Doing Away with Harm', *supra* note 4. On the second, see, among others, Victor Tadros, *Wrongs and Crimes* (Oxford: Oxford University Press, 2016), 175–185.

falling on me from the wind can harm me, just as someone dropping a tree on me can harm me. Second, non-agents can be harmed. If the tree falls onto my dog, Mouse, it harms her. Third, this account is morally neutral. When I justifiably defend myself against someone, I harm them regardless of my act being justified.

II. THE DESIRE EXPLANATION

Often, people desire not to be subjected to risk of harm. Desires might play some role in harm. Perhaps people are harmed through being at risk because their desires are frustrated. Perhaps Target is harmed because she has a desire not to be the subject of risk, a desire that Shooter frustrates. Call this *the Desire Explanation*.¹¹

There are several ways that the Desire Explanation might work, depending on the correct account of wellbeing.¹² Assume the Desire Theory of wellbeing, on which wellbeing consists only in the satisfaction of one's desires. In *Roulette*, if Target has a desire that Shooter not subject her to risk, Target is harmed by Shooter. This is because Shooter makes Target worse off than Target would have been, had she not played Russian roulette with Target. She makes Target worse off than Target would have been because she causes Target's desires to be more frustrated than they would have been, had she not acted as she did. This is one version of the Desire Explanation.

But one need not hold the Desire Theory of wellbeing to defend the Desire Explanation. There are other views of wellbeing in addition to the Desire Theory that see the frustration of one's desires as bad for an individual (though, unlike the Desire Theory, they do not imply that it is the only thing that is bad for an individual). For example, Hybrid Theories that subsume the Desire Theory will imply that the frustration of one's desires are bad for one. On these views, risk will be harmful in virtue of one's desires not to be at risk being frustrated. But let us assume the Desire Theory for ease of exposition.

¹¹ The Desire Explanation is defended by Finkelstein, 'Is Risk a Harm?', *supra* note 3. If the Desire Explanation succeeds, anyone who thinks desires are themselves relevant for wellbeing will be committed to the view that risk is itself harmful.

¹² It depends also on the answers to the questions posed in note 10.

Some try to force a dilemma on the defender of the Desire Explanation, both horns of which they take to be problematic.¹³ In subsection II.A, I argue this dilemma can be dissolved. In subsection II.B, I introduce a new problem that is not so easy to avoid.

A. Actual and Informed Desires

The Desire Theory says that what is good for an individual is only the satisfaction of their desires. We can ask whether the desires that are relevant for the Desire Theory are an individual's *actual* or *informed* desires.¹⁴ In *Cherry Pie*, suppose that I want some cherry pie and there is one in front of me. Unbeknown to me, I have developed a severe allergy to cherries.¹⁵ If the Desire Theory cares about my actual desires (call this the Actual Desire Theory), my life goes better, for me, if I get the pie than if I do not. This verdict seems odd. My life seems to go worse for me if I get the pie. (Below, we turn to a more sympathetic reading of this case.) More generally, as James Griffin writes, 'notoriously, we mistake our own interest'.¹⁶ In *Cherry Pie*, were I to be presented with all of the non-evaluative facts, I probably would not continue to desire a slice of pie. This might give us reason to think that the desires that are relevant for the Desire Theory are our *informed* desires. Call this the Informed Desire Theory.

I said that some people try to force a dilemma upon defenders of the Desire Explanation. On the first horn of the dilemma, assume the Actual Desire Theory. Stephen Perry suggests, if one holds the Actual Desire Theory, 'people can prefer or disprefer almost anything, so there would seem to be nothing in principle to prevent someone from preferring a risky to a non-risky state of affairs'.¹⁷ I take it that

¹³ Adler, 'Risk, Death and Harm', *supra* note 3 at 1251–1253; Stephen Perry, 'Risk, Harm, Interests, and Rights', in *Risk: Philosophical Perspectives*, ed. Tim Lewens (Chippenhams: Routledge, 2007), 200–201; Oberdiek, 'The Moral Significance of Risking', *supra* note 3 at 346–347.

¹⁴ E.g., Henry Sidgwick, *The Methods of Ethics*, 7th ed. (Indianapolis, IN: Hackett, 1981), 109–115; John Rawls, *A Theory of Justice*, Rev. ed. (Oxford: Oxford University Press, 1999), 365–372; James Griffin, *Well-Being: Its Meaning, Measurement and Moral Importance* (Oxford: Oxford University Press, 1986), 10–20. By actual desires, I mean something closer to an individual's *hypothetical* desires – the desire she would have, were someone to ask her. On the the move from *revealed* to *hypothetical* desires, see L. W. Sumner, *Welfare, Happiness, and Ethics* (Oxford: Clarendon, 1996), 119.

¹⁵ Chris Heathwood, 'Desire-Fulfillment Theory', in *The Routledge Handbook of Philosophy of Well-Being*, ed. Guy Fletcher (London: Routledge, 2016), 139.

¹⁶ Griffin, *Well-Being: Its Meaning, Measurement and Moral Importance*, *supra* note 14 at 10.

¹⁷ Perry, 'Risk, Harm, Interests, and Rights', *supra* note 13 at 200.

Perry's worry is meant to be that this means risk is not necessarily harmful because someone may not desire not to be at risk in some particular circumstances. John Oberdiek suggests the Actual Desire Theory of wellbeing is too implausible in general to even entertain whether risk is harmful on it. He writes, 'we often prefer, through ignorance or whatnot, what is in fact bad for us [and this] is decisive' against the Actual Desire Theory.¹⁸

On the second horn of the dilemma, assume the Informed Desire Theory. That theory says that what is good for an individual is having the desires satisfied that she would have were she in possession of all the non-evaluative facts. It is pressed, from this 'omniscient perspective' of fully informed desires, 'risk just disappears'.¹⁹ The idea is supposed to be, if one knows that some risk will materialise, there is no *risk* of harm, but harm in the straightforward sense; and if, on the contrary, one knows that the risk will *not* materialise, it will not be the case that one does not desire to be subjected to the putative risk. On either horn of the dilemma, then, the Desire Explanation fails.

The Desire Explanation can be defended on both fronts. Let us begin with the first horn of the dilemma. In response to Perry, suppose that someone is indifferent about being subjected to some particular risk. Perry is correct in thinking that this means that *that* risk is not harmful to that particular person in those particular circumstances, since there is no desire not to be subjected to risk that is frustrated. But this is not so much an objection to the Desire Explanation as statement of the Actual Desire Theory itself.²⁰ This fiat of the Desire Theory results from its *subjective* character as a theory of wellbeing: on the Desire Theory, 'getting a good life has to do with one's attitudes towards what one gets in life rather than the nature of those things themselves'.²¹ Many people endorse the Desire Theory *because of* its subjective character.²²

¹⁸ Oberdiek, 'The Moral Significance of Risking', *supra* note 3 at 347.

¹⁹ Oberdiek, 'The Moral Significance of Risking', *supra* note 3 at 346.

²⁰ It is also the case that one may desire a risky state of affairs over a non-risky state of affairs, fully informed, so Perry's objection applies equally to the Informed Desire Theory.

²¹ Heathwood, 'Desire-Fulfillment Theory', *supra* note 15 at 135.

²² For example, Railton says, 'what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware'. Peter Railton, 'Facts and Values', *Philosophical Topics* 14, no. 2 (1986): 9. (While the end of this quote indicates Railton endorses the *Informed* Desire Theory, the main point of the quote concerns the choice between *objective* and *subjective* theories of wellbeing.)

What of Oberdiek's rejection of the Actual Desire Theory because of its implausibility in reply to cases like *Cherry Pie*? We might not need to embrace the Informed Desire Theory to get the intuitively correct verdict in *Cherry Pie*. This is because the Actual Desire Theory is ambiguous between saying that it is good for me, all things considered, to eat the pie and that it is good for me, pro tanto, to eat the pie. If we embrace the second disambiguation, the Actual Desire Theory can maintain that it is good for me, pro tanto, to eat the cherry pie, but that it is not good for me all things considered. It might not be all-things-considered good for me because doing so will frustrate many of my other desires. In favour of this line of thinking, compare the following two states of affairs: (S1) I do not have a desire to eat the cherry pie, eat it nonetheless, and have an allergic reaction; (S2) I do have a desire to eat the cherry pie, eat it, and have an allergic reaction. My life goes worse for me in S1 than S2. The pro tanto disambiguation can explain this.²³

Given the preceding two paragraphs, the Actual Desire Theory horn of the dilemma does not seem as worrying as some have thought. Let us move on to consider the second horn of the dilemma. One might respond to the second horn of the dilemma by suggesting that there is a distinction between, on the one hand, being fully informed of all of the non-evaluative facts in relation to the circumstances of a decision before deciding, and, on the other hand, being fully informed in relation to the outcome of the decision. We might put the point in the following way: agents can be more (or fully) informed about their choices from an *ex ante* perspective (before the fact) without becoming fully informed about their choices from an *ex post* perspective (after the fact, with knowledge of whether the risk materialises). And, risk disappears in the way that Oberdiek suggests only on the second disambiguation of informed desires.²⁴

²³ For discussion, see Sumner, *Welfare, Happiness, and Ethics*, *supra* note 14 at 131; Chris Heathwood, 'The Problem of Defective Desires', *Australasian Journal of Philosophy* 83, no. 4 (2005): 487–504. The purpose of this paragraph has been to show that the Actual Desire Theory is not simply to be set aside. That said, I am sceptical of whether this pro tanto disambiguation will work in all cases. Suppose that I desire state of affairs S3 over S4 because I think that S4 will be painful and S3 not painful. I am mistaken. S3 is painful and S4 would not have been. What pro tanto desire is satisfied by my choosing S3?

²⁴ This will need to be slightly refined for one might think, if determinism is true *and* one knew all of the *ex ante* facts, they will be able to figure out all of the *ex post* facts. Thanks to Antony Duff and Sandra Marshall for this point. But we could go for a Reasonably Informed Ex Ante Desire Theory. Oberdiek does seem aware of this reply on behalf of the Desire Explanation: 'It may well be possible to craft an account of "sufficiently informed" preferences that have normative bite but which stop short of being fully informed, thus preserving some room for risk'. Oberdiek, 'The Moral Significance of Risking', *supra* note 3 at 347.

As an illustration, let us amend *Cherry Pie*. As above, I want a slice of cherry pie and there is one in front of me. Unbeknown to me, I have developed a severe allergy to cherries. Unlike above, suppose that this allergy results in an allergic reaction fifty percent of the time. When the allergy does not manifest, I have no adverse reaction. The Informed Desire Theory might say that what is good for me should be decided from the informed *ex ante* perspective (given knowledge of my allergy and the propensity of its occurrence) or from the *ex post* perspective (given both knowledge of my allergy and knowledge of whether, in this instance, I will have an allergic reaction to the pie). Risk does not disappear on the *ex ante* perspective of the Informed Desire Theory.

We have been considering the following dilemma that threatens the Desire Explanation. If one endorses the Actual Desire Theory, agents may desire a risky state of affairs over a non-risky state of affairs and so risk will not necessarily be harmful. However, this is just a fiat of the Desire Theory's subjective character as a theory of wellbeing. And if one endorses the Informed Desire Theory, one might think that risk just disappears insofar as the agent will be informed of whether or not the risk will manifest itself, and so would only desire not to be subject to risks that will materialise. While this might be true on the Informed Desire Theory (Ex Post Perspective), this is not true on the Informed Desire Theory (Ex Ante Perspective). Risk might itself be harmful on the Actual Desire Theory or on the Informed Desire Theory (Ex Ante Perspective).

B. *Derivative and Non-Derivative Desires*

We are trying to determine whether risk of harm is itself harmful. Since it might be bad for us when our desires are frustrated, and we have desires not to be at risk of harm, risk might be harmful when our desires not to be at risk are frustrated. This is the Desire Explanation.

Recall again that the Desire Theory (of wellbeing) says that what is good for an individual is the fulfilment of their desires. There are several choice-points for the Desire Theorist. Above we compared the Actual and Informed Desire Theories. The choice-point that is important for our purposes in this subsection is whether the desires

that are relevant are both individuals' *derivative* and *non-derivative* desires or only their *non-derivative* desires.²⁵ Let us say a person's desire for something is non-derivative if she desires it for its own sake, and her desire for something is derivative if she desires it, but not for its own sake. On what we can call the Non-Derivative Desire Theory, what is good for an individual is only the fulfilment of their non-derivative desires. On the Unrestricted Desire Theory, what is good for an individual is both the fulfilment of their derivative and non-derivative desires.

Suppose that Dan wants it to snow in the mountains so that he can get some good skiing in. It does snow but Dan does not end up going skiing.²⁶ Does the satisfaction of Dan's desire that it snow in the mountains make his life go better for him, even though he does not end up going skiing? Dan's desire that it snow was only a derivative desire. The satisfaction of his desire that it snow seems to derive its value only from his desire to get some good skiing in. Dan's desire that he get some good skiing in is a non-derivative desire (or at the least, closer to a non-derivative desire). The Non-Derivative Desire Theory says Dan's life does not go better for him. The Unrestricted Desire Theory says his life does go better for him.

We have good reason to endorse the Non-Derivative Desire Theory. First, the Unrestricted Desire Theory is intuitively implausible. As Chris Heathwood suggests, 'intuitively, the fulfillment of [Dan's] desire that it snow was not in the end of any benefit to [Dan]'.²⁷ Second, compare the following two worlds.²⁸ In both worlds, I want to eat an apple. In both worlds, one apple remains on a tree. In world 1, I need only to reach out to get the apple. In world 2, I need a ladder to reach the apple. To satisfy my desire to get the apple, I will also desire a ladder, though this desire will be only derivative. Luckily, there will be a ladder nearby. If the Unrestricted Desire Theory is correct, my life goes better for me in world 2 than it does in world 1. This is because, in both worlds, I get to satisfy my

²⁵ Sidgwick, *The Methods of Ethics*, *supra* note 14 at 109; Heathwood, 'The Problem of Defective Desires', *supra* note 23; Heathwood, 'Desire-Fulfillment Theory', *supra* note 15 at 139. For relevant discussion concerning the Desire Theory of reasons, see: Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), 117; Derek Parfit, *On What Matters: Volume One* (Oxford: Oxford University Press, 2011), 58–59.

²⁶ Heathwood, 'Desire-Fulfillment Theory', *supra* note 15 at 139.

²⁷ *Id.*

²⁸ The case is inspired by Parfit, *On What Matters: Volume One*, *supra* note 25 at 59.

desire to eat an apple. But in world 2, I get to satisfy an *additional* desire – my desire for a ladder. But it is implausible that my life goes better for me in world 2 than world 1 in virtue of the satisfaction of the additional derivative desire.

So, we have good reason to prefer the Non-Derivative Desire Theory. Returning to risk, the Desire Explanation says Shooter harms Target in virtue of frustrating Target's desire not to be at risk. However, given our discussion of derivative and non-derivative desires, we now know this follows only if Target's desire is non-derivative. But Target's desire that she not be at risk seems to be a derivative desire – it is derivative on, at the least, her non-derivative desire that she not be harmed. So, Target is not made worse off by the frustration of her derivative desire not to be at risk. Contrary to the Desire Explanation, we cannot hold that risk is harmful insofar as people desire not to be subjected to risk.

The preceding argument rests on the claim that Target's desire that Shooter not subject her to risk is merely a derivative desire. This seems fairly obvious to me – why would one desire not to be at risk of harm *for its own sake*? But perhaps it is not as obvious to everyone. To respond to this point, let me offer an *argument in the alternative*. On the one hand, I bolster support for the view that desires not to be at risk of harm are merely derivative. On the other hand, I show it does not matter if I am mistaken, and some people do desire not to be at risk for non-derivative reasons.

Suppose that Bloggs is indifferent between being harmed to some degree on some occasion, and not being harmed. Suppose that her indifference is rational, for example, because she would be compensated were she to be harmed to the level that would leave her indifferent between having not been harmed, and having been harmed and then compensated. Now, despite this indifference, suppose that Bloggs desires that she not be subject to risk of harm. What we have is a case in which Bloggs holds the desire not to be at risk of harm, though not the desire not to be harmed. Once we factor away any non-derivative reasons for which she might hold this desire not to be at risk, the desire looks very odd to me. (It is not that Bloggs desires not to be at risk because being so will make her anxious or on edge, nor that Bloggs is risk-adverse, for both of these features will already be accommodated when determining the

compensation owed.) The explanation for why this desire looks odd, I submit, is that desires not to be at risk are ordinarily merely derivative on desires not to be harmed – but we have specified in this case that there is no non-derivative desire not to be harmed for the derivative desire not to be at risk of harm to latch onto. And since, I take it, we do not think that many of our desires not to be at risk are odd in this way, this suggests that many of our desires not to be at risk are merely derivative.

Yet, one might object that this does not show that non-derivative desires not to be at risk are not possible, nor does it show even that they are rare. Perhaps people sometimes do desire not to be at risk for non-derivative reasons. And because of this, risk *can* sometimes be itself harmful given the Desire Explanation – namely, when someone has a non-derivative desire not to be at risk of harm. It is worth noting a strange asymmetry this leaves the defender of the Desire Explanation maintaining: while risk is itself harmful to those who desire not to be subjected to risk for non-derivative reasons, risk is not itself harmful to those who desire not to be subjected to risk for merely derivative reasons (for example, because they do not want to be harmed). I find this asymmetry very counterintuitive. But, regardless, even if the objection under consideration in this paragraph is correct – even if risk is harmful when people desire not to be at risk for non-derivative reasons – recall this paper's first question: Assuming risk is of moral significance, does risk's moral significance lay in its being harmful? And our answer has to be no. If risk's moral significance did lay in its being harmful, this would imply that pure risk is not of moral significance when imposed on those who desire not to be subjected to risk for merely derivative reasons.²⁹ But that seems false: what Shooter does to Target is of moral significance, even if Target's desire not to be at risk is merely derivative.

I have argued that the Desire Explanation fails to explain why risk would be harmful. The desire that we not be at risk of harm is, most plausibly, a derivative desire (derivative on our non-derivative desire not to be harmed). The frustration of a derivative desire does not *itself* constitute a diminishment to one's wellbeing and, thereby, a

²⁹ Perhaps one might reply by saying we ought to endorse the Ideal Desire Theory and, under idealised conditions, people *should* or *would* desire not to be subjected to risk for non-derivative reasons. Yet, this begs the question. Further, if we want the desire not to be at risk to ground risk's moral significance, it cannot be that we desire not to be at risk *because* we think it is of moral significance.

harm. And even if it is possible to desire not to be subject to risk for non-derivative reasons, this means that risk's moral significance cannot lay in its being harmful, for many desires not to be at risk will not be like this, and yet those risks will still be of moral significance.

III. THE AUTONOMY EXPLANATION

A. *Autonomy and Harm*

There is another way that risk might be harmful. One reason we might think the satisfaction of Dan's desire that it snows in the mountains does promote his wellbeing is that it gives him the *option* of getting some good skiing in. And this is true even if he does not end up going skiing. More generally, we might think having the option to do things is valuable to us, even if we do not end up doing those things.

One reason why one may think this is if one believes having options is partly constitutive of autonomy. In particular, one might endorse:

Adequate Range. For an individual's choice to be autonomous, she must have an adequate range of valuable options.³⁰

Oberdiek motivates Adequate Range by saying that 'one is autonomous when one can plot and pursue one's own worthwhile path, and to do this, one needs to have access to a range of valuable options'.³¹ We return to Adequate Range below (section III.B.1).

Returning to risk, suppose that autonomy plays some role in wellbeing – a person's life goes better or worse, for them, when their autonomy is promoted or frustrated. Now consider

Two Doors. Chooser is faced with a choice. Let choosing between two doors, *A* and *B*, stand in for an autonomous choice between two different valuable things. Unbeknown to Chooser, Locker locks door *B* (he stops her from being able to do whatever valuable thing door *B* takes the place of). Chooser chooses door *A*.

Locker risks harm to Chooser. Straightforwardly, he risks it being the case that Chooser chooses door *B* and is unable to do whatever valuable thing door *B* takes the place of. If Adequate Range is correct, Locker frustrates Chooser's autonomy, even if Chooser is unaware of this. Locker frustrates Chooser's autonomy because he stops her from having an adequate range of valuable options to

³⁰ Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), 372.

³¹ Oberdiek, *Imposing Risk*, *supra* note 3 at 9.

choose from. Because he frustrates Chooser's autonomy and because, we are assuming, Chooser's autonomy is partly constitutive of her wellbeing, he harms Chooser. More generally,

if *Y* subjects *X* to risk and, thereby, frustrates *X*'s autonomy, *Y* harms *X*.

Let us call this the *Autonomy Explanation*. Given the Autonomy Explanation, risk is – sometimes – harmful.³²

The Autonomy Explanation has support in the literature. Oberdiek suggests that, 'while imposing risk does not involve material harm [...] it can nevertheless constitute a setback to a nonmaterial autonomy interest of a certain kind'.³³ Subjecting someone to risk

effectively attaches sanctions to or normatively forecloses certain options that would otherwise be available to the individual, thereby narrowing the risked person's set of worthwhile opportunities. Narrowing one's open future diminishes one's autonomy suitably understood, and it is in this that the moral significance and thus the potential impermissibility of pure risking lies.³⁴

Oberdiek offers us the following analogy. By laying a trap, though I may not experientially affect anyone, I do nonexperientially affect people: 'This is because the trap takes away the option, or more accurately renders unacceptable the exercise of the option, of stepping where the trap has been set'.³⁵ If enough traps were laid, one's autonomy would be completely frustrated.

Seth Lazar holds a somewhat similar view. He says that, 'if others avoidably make us dependant on [luck] for our avoidance of wrongful harm, they harm us' and that, when one subjects another to risk, this 'contravenes [an] important interest [...] in *being secure*'.³⁶ He thinks this for at least two reasons: first, being secure serves 'contingent benefits such as peace of mind, or the ability to plan for the future'; second, 'the more you depend on luck, the less control you have over your life, and so the less autonomy you have'.³⁷ The first of these reasons does not concern pure risk, so can be set aside.

³² Ben Colburn has pointed out that it is possible that one can remove options from others but not make it the case that Adequate Range is not satisfied. This is possible if the individual whose options have been limited still has *an* adequate range of sufficiently valuable and diverse options. In cases of this sort, exposing others to risk will not be harmful. The defender of the Autonomy Solution could reply by revising Adequate Range and saying that, whenever one's options are diminished, one's autonomy is frustrated to *some* extent.

³³ Oberdiek, 'The Moral Significance of Risking', *supra* note 3 at 342.

³⁴ *Id.* at 351–352.

³⁵ *Id.* at 352.

³⁶ Lazar, 'Risky Killing and the Ethics of War', *supra* note 3 at 102; Lazar, 'Risky Killing', *supra* note 3 at 7.

³⁷ Lazar, 'Risky Killing', *supra* note 3 at 8; Lazar, 'Risky Killing and the Ethics of War', *supra* note 3 at 102. See subsection IV for a third reason Lazar offers.

The second reason appeals to autonomy, and so makes Lazar vulnerable to the objections raised against the Autonomy Explanation below.

Maria Ferretti holds another similar view. She thinks, 'what is specifically morally problematic about risk consists in not 'respecting people as moral agents', which she cashes out in terms of a diminishment of overall freedom.³⁸ While this account is not directly answering the questions of whether risk is harmful, it does fall foul of problems similar to those discussed below (subsections III.B.2–III.B.3).

B. Against the Autonomy Explanation

This subsection raises three objections against the Autonomy Explanation.

1. Theoretical Commitments

The Autonomy Explanation relies on a substantive, controversial theoretical commitment concerning autonomy: namely, one needs to endorse Adequate Range.³⁹ Adequate Range says that, for one to enjoy an autonomous choice, one needs to have chosen from an adequate range of valuable options. But some views of autonomy deny this. And they may have good reason to, for example, since Frankfurt cases put pressure on Adequate Range. Instead, on these views that deny Adequate Range, whether a choice is autonomous depends only on how the decision was arrived at. Let me briefly present a view of this sort, which should serve for illustrative purposes.

Serena Olsaretti thinks that a choice is *not* voluntary 'if it was made because no other acceptable alternative was available'.⁴⁰ Olsaretti continues that we should distinguish between 'first-order desires and wishes, which is what we focus on when considering the

³⁸ Maria P Ferretti, 'Risk Imposition and Freedom', *Politics, Philosophy & Economics* 15, no. 3 (2016): 262.

³⁹ One needs also to think that autonomy is partly constitutive of wellbeing. For example, suppose one is a Hedonist – one thinks wellbeing consists only of pleasure. On this view, frustrating someone's autonomy will harm them only when you also make them worse off, in terms of their wellbeing, than they would have been were you not to have frustrated their autonomy. But since we are examining pure risk, in which the risked harm does not materialise, risk is not itself harmful. Thanks to two anonymous reviewers for discussion on this.

⁴⁰ Serena Olsaretti, 'Freedom, Force and Choice: Against the Rights-Based Definition of Voluntariness', *Journal of Political Philosophy* 6, no. 1 (1998): 54.

voluntariness of actions, and autonomy as the second-order capacity to reflect critically over one's first-order preferences and desires, and to decide which ones to act on'.⁴¹ Building off Olsaretti's view of voluntariness, her view of autonomy would say that a choice is autonomous only if, first, it was not made because no other acceptable alternative was available and, second, because it was consistent with a second-order desire (that itself does not exist because no other acceptable alternatives are available).

On this view of autonomy, one does not actually *need* an adequate range of valuable options to choose from. Rather, one needs, first, not to have chosen because no other acceptable option was available and, second, one's choice needs to be consistent with one's second-order desires. While lacking an adequate range of valuable options will often stop one from making a choice because no other option was available, this is merely contingent. And because we are considering *pure risk*, where one is unaware that one's options are limited, one's autonomy will not be frustrated in this way. For example, recall *Two Doors*. We can suppose Chooser thinks both options, doors A and B, are valuable and available to her. She is mistaken, since Locker has locked door B. However, because she chooses door A and does not know door B is unavailable to her, Olsaretti's view will allow that choice to be autonomous, since Chooser did not choose it because no other option was available. Oberdiek's view, in contrast, will say her choice is not autonomous, since she did not in fact choose from an adequate range of options.

The objection raised in this subsection is limited in scope, for I have taken no stand on whether to endorse Adequate Range. But it raises two points. First, the Autonomy Explanation can be appealed to only by those who accept Adequate Range. Second, even if one endorses Adequate Range, one might think that, other things being equal, an answer to the question 'Is risk harmful?' is unattractive to the extent that it relies on controversial normative commitments – and the Autonomy Explanation relies on one such commitment.

2. *Non-Autonomous Individuals*

The Autonomy Explanation implies that risk is harmful only for people who are capable, at that time, of leading autonomous lives.

⁴¹ *Id.* at 73.

This is because it is only those people who can have their autonomy frustrated *in virtue of* having valuable options removed.⁴² This precludes, among others, people with undeveloped, compromised, or damaged rational capacities (for example, very young children, the severely mentally disabled, and those suffering from the later stages of Alzheimer's disease).⁴³ We can draw two conclusions from this. First, the Autonomy Explanation does not say that risk is harmful for those without autonomy. If the moral significance of risk lies in risk's being harmful, this means that exposing those without a capacity for autonomy to pure risk is not morally significant. But this is clearly false. Shooter's playing Russian roulette against people in these categories is clearly itself morally significant.

Second, we can posit the following argument. The Autonomy Explanation leaves us with an asymmetry: exposing autonomous individuals to risk is itself harmful, whereas exposing people with undeveloped, compromised, or damaged rational capacities to risk is not itself harmful. This asymmetry *itself* is intuitively implausible. This suggests the Autonomy Explanation is implausible.

Let me say a little more about why the Autonomy Explanation implies risk is not itself harmful for those without a capacity for autonomy. Suppose that Villain kills someone with severe mental disabilities. Villain does not frustrate this person's autonomy because they have no autonomy for the Villain to frustrate. *A fortiori*, Villain does not frustrate this person's autonomy by exposing them to *risk* of death. So, subjecting this person to risk is not itself harmful (given the Autonomy Explanation).

The same holds if Villain subjects a baby to risk of death. If Villain were to kill the baby, we would not say she frustrates the baby's autonomy (for, again, the baby has no autonomy for Villain to frustrate) – and we certainly would not say that she frustrates the baby's autonomy in virtue of denying her a valuable range of options to choose from. So, we should not say Villain frustrates the baby's autonomy by exposing her to risk of death. What the Villain may do were she to kill the baby is deny the baby the *opportunity* of having

⁴² See, also, Placani, 'When the Risk of Harm Harms', *supra* note 3 at 90; Madeleine Hayenhjelm, 'Oberdieck, John. Imposing Risk: A Normative Framework', *Ethics* 129, no. 3 (2019): 496.

⁴³ It also includes non-human animals. But my dog can be harmed by kicking her just as I can be harmed by kicking me. The Autonomy Solution implies, while subjecting me to risk is itself harmful, subjecting my dog to risk is *not* harmful.

an autonomous life. However, if the risked harm does not materialise and the baby lives, Villain does not deny the baby this opportunity. So, she has not interfered with the autonomous life the baby will *later* come to have.

3. *Roulette and Autonomy*

The Autonomy Explanation works by showing that, when one subjects another to risk, one might frustrate the other's autonomy by stopping them from having an adequate range of valuable options to choose from, independently of whether the risked harm materialises. In *Two Doors*, Locker frustrates Chooser's autonomy by removing the option of going through door *B*. Thereby, he stops Chooser from having an adequate range of valuable options. But this means that the Autonomy Explanation applies only to cases in which the risk affects the exercise of a potentially autonomous choice – and it applies only in virtue of stopping *that* choice from being autonomous. This analysis does not straightforwardly extend to *Roulette*. Target is asleep. She is not exercising any autonomous choices at the time at which the risk is imposed. And so, it is not obvious that Shooter's subjecting her to risk removes any valuable options from her in a way that undermines her autonomy. We can ask, what valuable option does Shooter remove? And even if we can provide an answer to this question, will that answer capture the moral significance of Shooter's act in *Roulette*?

Shooter risks it being the case that Target does not have 'any future choice', autonomous or otherwise. But that itself does not *frustrate* Target's autonomy – it merely *risks* frustrating her autonomy. To see this, consider the following case.

Two Doors (Two Choices). Chooser is faced with a choice. Let choosing between two doors, *A* and *B*, stand in for a potentially autonomous choice. After going through either door, Chooser will then be faced with another choice. Let choosing between A_1 and A_2 stand in for a potentially autonomous choice that she would face if she went through door *A*. Let choosing between B_1 and B_2 stand in for a potentially autonomous choice that she would face if she went through door *B*. Unbeknown to Chooser, Locker locks door *B*. Chooser chooses door *A*.

Chooser is now choosing between A_1 and A_2 . Is *that* choice autonomous? Yes. Before going through the door, Chooser faced two potentially autonomous choices: Choice 1, between doors *A* and *B*, and Choice 2, between A_1 and A_2 or between B_1 and B_2 . If Adequate Range is correct, Locker stops Choice 1 from being autonomous – he stops Chooser from having an adequate range of

valuable options. However, he does not stop Choice 2 from being autonomous. He merely *risks* Chooser not being able to make that Choice. Similarly, in *Roulette*, though Shooter might, for example, stop Target's choice of going to sleep from being autonomous, she does not stop any of Target's future choices from being autonomous.

We can further support the verdict that Choice 2 was autonomous, despite Choice 1 not being autonomous. If Choice 2 is rendered non-autonomous in virtue of Choice 1 not being autonomous, this implies a single non-autonomous choice can taint all future choices, rendering them non-autonomous. But this is implausible.

C. *Moving Beyond Autonomy*

The previous objection leads to a more general one. The Autonomy Explanation says that risk is harmful *because*, when one is exposed to risk, one's autonomy is frustrated. Accepting for the sake of argument that risk is harmful, is autonomy the correct sort of explanation *why* risk would itself be harmful? I am not convinced.⁴⁴

Let us conclude our discussion of the Autonomy Explanation. The Autonomy Explanation can be appealed to only by those who accept Adequate Range. It leaves us with several implausible asymmetries – it is odd that exposing adults with autonomy to risk is (in some circumstances) harmful, while exposing individuals with undeveloped, compromised, or damaged rational capacities to risk is not harmful. It is odd that Locker's exposing Chooser to risk is harmful in *Two Doors* but that Shooter's exposing Target to risk in *Roulette* is not harmful. I have also suggested that explaining why risk is itself harmful by reference to autonomy does not seem wholly satisfying. All of this gives us good reason to think that the Autonomy Explanation fails to show that risk is harmful.

Some might object to the preceding. First, they say, 'While you are undecided on our controversial theoretical commitment, I think that Adequate Range is correct.' Second, they might argue that the second and third objections do not impugn the Autonomy Explanation as an answer to the question 'Is risk itself harmful?' because

⁴⁴ I take it Placani is making a similar point when she says, 'At least when it comes to such egregious acts [as *Roulette*], Oberdiek's position seems a bit odd'. Placani, 'When the Risk of Harm Harms', *supra* note 3 at 90.

these objections show only that there are *some* cases in which risk is not harmful. But this does not show that risk is not harmful in any case. For those who argue in this way, we can note a conservative conclusion: the moral significance and potential impermissibility of pure risk imposition cannot lie in the Autonomy Explanation. It is *too* restrictive.⁴⁵

IV. DISRESPECT, DISPOSITIONS, AND PLANNING

I have argued the Desire and Autonomy Explanations have serious limitations and so, even if risk is sometimes harmful, on either view risk's moral significance cannot lay in its being harmful. But perhaps there might be other ways to claim that risk is harmful. In this section, I briefly object to three additional views on which risk is itself harmful. Then, in the following section, I raise a general worry with the very idea that risk is harmful.

Adriana Placani thinks that, in cases in which risk undermines people's moral status as agents, risk is harmful in virtue of setting back their dignity-interests.⁴⁶ Recall the first question that concerns this paper: 'Assuming risk is of moral significance, does risk's moral significance lay in its being harmful?' I think Placani's account gets things backwards as it concerns this question – subjecting others to risk is disrespectful *because* we are under directed duties not to subject others to risk of harm, so appealing to disrespect is not going to help us ground risk's moral significance.⁴⁷

There are additional problems. Placani is quite explicit that she is limiting her scope to 'risks imposed by one agent or multiple agents onto others [and] only consider[s] those risks of harm imposed intentionally onto others and to which no consent has been given'.⁴⁸ (As we see, Placani should say she limits herself to *wrongful* risks.) Limiting her focus in this way leads to some strange results. First, it can accommodate risk being harmful *only* when imposed by other agents, because only other agents can disrespect us. However, things

⁴⁵ Recall, Oberdiek writes: '[n]arrowing one's open future diminishes one's autonomy suitably understood, and it is *in this* that *the* moral significance and *the* potential impermissibility of pure risking lies' (my emphasis). 'The Moral Significance of Risking', *supra* note 3 at 351–352.

⁴⁶ Placani, 'When the Risk of Harm Harms', *supra* note 3.

⁴⁷ See Kagan for a similar argument about the appeal to disrespect in moral theory: Shelly Kagan, *The Limits of Morality* (Oxford: Oxford University Press, 1989), 176.

⁴⁸ Placani, 'When the Risk of Harm Harms', *supra* note 3 at 80.

other than agents can harm us, such as trees and animals. This leaves strange asymmetries: as suggested in the introduction, a tree's falling on me from the wind can harm me, just as someone dropping a tree on me can harm me, but risk of the tree falling on me is harmful only if someone subjects me to that risk. Second, suppose I subject someone to risk of harm. Whether this is harmful will depend, on Placani's account, on whether I am acting unjustifiably. Yet, the materialisation of the risked harm is harmful regardless of whether I act unjustifiably. This is because we do not disrespect others when we act justifiably towards them, but we do harm them nonetheless. For example, I do not disrespect a culpable aggressor when I justifiably harm them in self-defence, but I do harm them. This again leaves a strange asymmetry: unjustified risk of harm is itself harmful, but justified risk of harm is not harmful.

Let us turn to a second view. In section III, I said that Lazar says that when one subjects another to risk, this 'contravenes [an] important interest [...] in *being secure*'.⁴⁹ We looked at two reasons he endorses this view above (that it allows us to plan and enjoy autonomous choices). A third reason he offers is that 'one's security is often grounded in others' positive dispositions towards one', and 'I am better off just by [others] having this positive disposition towards me'.⁵⁰ But it is unclear how others' dispositions *in themselves* affect my wellbeing, unless I have valuable relationships with them. Setting this worry aside, like Placani's view, this reading of Lazar's view can accommodate risk being harmful *only* when imposed by other agents.

Finally, third, Stephen John thinks risk is harmful because it undermines agents' capacity to form reasonable plans.⁵¹ I agree with Jonathan Herington that one needs only to *believe* one is free from risk for one to enjoy the good of planning.⁵² But one's beliefs of this sort are not affected in cases of pure risk. So, assuming that pure risk is of moral significance, risk's moral significance cannot lay in its being harmful. Further, John's focus on planning leaves his view vulnerable to objections raised against the Autonomy Explanation

⁴⁹ Lazar, 'Risky Killing', *supra* note 3 at 7.

⁵⁰ Lazar, 'Risky Killing', *supra* note 3 at 8–9.

⁵¹ Stephen John, 'Security, Knowledge and Well-Being', *Journal of Moral Philosophy* 8, no. 1 (2011): 68–91.

⁵² Jonathan Herington, 'The Contribution of Security to Well-Being', *Journal of Ethics and Social Philosophy* 14, no. 3 (2018): 194–198.

because, for example, babies do not enjoy the good of planning, so cannot be harmed in virtue of frustrating that good.

V. THE MAGNITUDE OF HARM PROBLEM

We have now worked through several explanations of why risk might be harmful. Let us turn to a final argument against the view that risk is harmful, whichever explanation one offers. Commonly, I take it that we do not think that risk of harm is itself harmful. That much should be indicated by my saying ‘*risk of harm*’ and not ‘*risk of more harm*’. And while this itself is perhaps not a decisive objection, it does lead to a problem with the very idea that risk is harmful.

Suppose that Threatener imposes a 0.5 risk of harm, h , onto Victim₁, and a 0.25 risk of the same harm, h , onto Victim₂ (for example, Threatener risks electrocuting each victim). All else is equal. It seems clear that if risk is harmful, the greater the probability of some harm occurring, the more harmful being exposed to the risk is. This implies, if the risks do not materialise for either victim, Victim₁ is harmed more than Victim₂. Assuming that risk is harmful, this is intuitively plausible. But if risk is itself harmful then, were both risked harms to materialise, this implies Threatener harms Victim₁ more than she harms Victim₂. This is because both victims suffer the same harm (the electrocution) and Victim₁ also suffers a greater additional harm having been at a greater risk than Victim₂. This is intuitively implausible. Were one to be presented with the original case and told that both risks materialise, we would want to say that each victim has been harmed to the same extent. Call this the Magnitude of Harm Problem.⁵³ (Note, the Magnitude of Harm Problem is consistent with the verdict that Threatener *wrongs* Victim₁ to a greater extent than she *wrongs* Victim₂. Being able to say this simply requires offering an account of the moral significance of risk that does not rely on risk’s being harmful.)

Put differently, we can distinguish between the *risk harm* (the harmfulness of being exposed to risk), the *outcome harm* (the harm-

⁵³ The Magnitude of Harm Problem relies on an intuition – that Victim₁ and Victim₂ are harmed to the same extent. Both Finkelstein and Placani think risk of harm is itself harmful and have this intuition in cases similar to mine: Finkelstein, ‘Is Risk a Harm?’, *supra* note 3 at 990–901; Placani, ‘When the Risk of Harm Harms’, *supra* note 3 at 96. One way to move beyond the intuition is to think about the problem’s implications. If risk of harm is itself harmful then, given that we tend to think the level of compensation that is owed is proportionate to the harm caused, this would imply that Victim₁ is owed more in compensation than Victim₂. But that is odd. See, also, the following note.

fulness of the materialisation of the risk), and the *all-things-considered* harm. Since the risk harm is greater for Victim₁ than Victim₂, and the outcome harm is the same for each victim, the all-things-considered harm must be greater for Victim₁ than Victim₂. But this seems implausible.

In support of the Magnitude of Harm Problem, recall from the introduction how the Counterfactual Account of Harm allows events caused by non-agents to be harmful. A tree falling in the wind can harm me just as much as a person's dropping a tree on me can. The Magnitude of Harm Problem seems even more counterintuitive when we think of non-agential harms. Suppose Victim₁ and Victim₂ are at different risks of being harmed by some natural event, for example, being struck by lightning. Now suppose both victims are struck. It is very implausible that Victim₁ is harmed more than Victim₂ insofar as she was subjected to greater risk than Victim₂.⁵⁴

One reply to the Magnitude of Harm Problem is that, when we say Victim₁ and Victim₂ suffer the same harm, we implicitly disambiguate the harms they suffer. First, we have the *different* risk harms they suffer through being subjected to different risks of the same harm. Second, we have the *same* outcome harm that they suffer from the electrocution. When we say that both victims are harmed to the same extent, we are referring to the same outcome harm that they suffer from the electrocution. We are not referring to the different risk harms they suffer.

This reply will not do. If the oddness of thinking that Victim₁ and Victim₂ suffer different harms was reducible to an ambiguity relating to which harm we were asking after, the counterintuitiveness that the Magnitude of Harm Problem identifies would vanish when we re-ask,

all-things-considered, is Victim₁ harmed more than Victim₂?

But the answer to this *unambiguous* question seems to be: 'No, both victims are harmed equally'. The counterintuitiveness of saying that the victims are harmed to a different extent cannot be explained away by appealing to ambiguity.

⁵⁴ Here is another version of the Magnitude of Harm Problem. Suppose that Victim₃ is exposed to probability, x , of harm h ; Victim₄ is exposed to a greater probability, y , of a smaller harm, g . If both risks materialise, there is a some set of values for our four variables that mean that the victims have been harmed to the same extent. But, again, this is implausible.

Claire Finkelstein offers us a potential solution to the Magnitude of Harm Problem. She says that ‘the disvalue of the risk harm is absorbed into the loss in welfare if the risk actually materializes’.⁵⁵ So, in reply to our case above, if neither of Threatener’s risks materialise, both victims are harmed merely by being subject to risk. Victim₁ is harmed more than Victim₂. But if both risks materialise, the different risk harms they suffer are absorbed into the outcome harms. And the outcome harms are equal, so the all-things-considered harms are equal. Call this the Absorption Solution.

There are at least two problems with the Absorption Solution. First, it is very difficult to see what principled explanation might lie behind the Absorption Solution. *Why* does risk harm persist only when the risked harm does not materialise?⁵⁶

Second, I am very unsure of how the Absorption Solution is supposed to work. Suppose the outcome harm of either victim being electrocuted is 100. And suppose, for illustrative purposes, we apply a crude form of expected utility theory to determine how harmful the risk of being electrocuted is. Victim₁ has a 0.5 chance of being electrocuted: she is harmed by 50 through being subjected to risk of being electrocuted. Victim₂ has a 0.25 chance of being electrocuted: she is harmed by 25 through being subjected to risk of being electrocuted. The Magnitude of Harm Problem says that, if both risks materialise, the all-things-considered harm for Victim₁ is 150 and for Victim₂ is 125. This is implausible.

The Absorption Solution says the all-things-considered harms are equal because the risk harm is absorbed into the outcome harm, and the outcome harms are equal. Yet how can this be? If the risked harms do not materialise, Victim₁ is harmed by 50 and Victim₂ is harmed by 25. If the risked harms do materialise, both victims are harmed by 100, all things considered. But this means the extent to which Victim₁ is worse off if the risked harm materialises than if it

⁵⁵ Finkelstein is responding to two slightly different problems: Finkelstein, ‘Is Risk a Harm?’, *supra* note 3 at 993. She thinks we should distinguish between intentionally harmful acts and risky harmful acts. This dichotomy is inconsistent with the idea that individuals can intentionally subject others to risk of harm. But this is clearly possible. Finkelstein’s solution is also endorsed by Placani, ‘When the Risk of Harm Harms’, *supra* note 3 at 96.

⁵⁶ One might reply, in cases in which the risked harm materialises, this shows the risk of harm is actually 1. So, both Victim₁ and Victim₂ are harmed by the same amount. Yet, on what basis do we get to say that the probability of harm is actually 1? I do not peruse this question here because whatever we say will imply, on grounds of parity, that if the risked harm does not materialise, the risk was actually 0; and so, the risk harm will be nothing.

does not materialise is 50 and the extent to which Victim₂ is worse off if the risked harm materialises than if it does not materialise is 75. The extent to which either victim is worse off if the risked harm materialises than if it does not *just is* the outcome harm. So, if the Absorption Solution is correct, the outcome harms cannot be equal – the outcome harm for Victim₁ is 50 and for Victim₂ 75. But this is implausible. How can the outcome harms be different depending on the level of risk one was subjected to?

Perhaps one might suggest that this is the wrong way to think about the Absorption Solution, and that the outcome harms are 100 for each victim, even though Victim₁ is made worse off by 50 and Victim₂ is made worse off by 75 if the risked harms materialise. One might explain this by saying the risk harm *pre-empts* some of the outcome harm. Consider the following standard example of pre-empted harm. Suppose Ann shoots Beth. Had Ann not shot Beth, Anna would have shot Beth. Intuitively, Ann harms Beth, even though Beth is no worse off than she would have been had Ann not shot her (she is no worse off than she would have been since Anna would have shot her). We might think the outcome harm is like Ann's shooting Beth and the risk harm is like Anna's shooting Beth.

Now, it is not obvious what, if anything, defenders of the Counterfactual Account of Harm should say to deal with pre-empted harm to accommodate the idea that one is harmed when that harm is pre-empted, even though it leaves one no worse off than one would have been.⁵⁷ In any case, if a defender of the Absorption Solution were to go this way, they would owe us a compelling explanation of *why* risk harm pre-empts some of the harmfulness of the outcome harm. That is, we are owed an explanation of why, even though risk is *itself* harmful, its badness is pre-empted by the badness of the risked harm materialising. And I do not think a good answer is forthcoming.

Here is what one answer could look like. Suppose one thinks that risk is itself harmful because one endorses the Desire Explanation. Then one could say, 'Well, because one only has a desire not to be subject to risk of harm *because* one does not want to be harmed, the desire not to be at risk of harm does not itself matter when one ends up being harmed'. Yet it is for this reason, I argued, the Desire

⁵⁷ See, e.g., Bradley, 'Doing Away with Harm', *supra* note 4; Neil Feit, 'Plural Harm', *Philosophy and Phenomenological Research* 90, no. 2 (2015): 361–388; Tadros, *Wrongs and Crimes*, *supra* note 10.

Explanation fails when one's desire not to be at risk is merely derivative. The frustration of merely derivative desires is not itself bad for us, so risk would not itself be harmful. Now, I did note that setting back a non-derivative desire not to be at risk might itself matter for one's wellbeing. Yet, if one's desire not to be at risk is non-derivative, I am left unsure why it would be subsumed by, or pre-empted by, the frustration of one's desire not to be harmed – since one does not have that desire not to be at risk of harm *only because* one desires not to be harmed. The task, then, is to offer a compelling explanation of why the harmfulness of being subjected to risk of harm is pre-empted by the harmfulness of the risked harm materialising, which does not undercut the view that risk is itself harmful.

Another explanation one could offer builds on Placani's view that risk is sometimes harmful because it sets back agents' dignity-interests. She could say that it is not as if another dignity-interest is setback if the risked harm materialises. Rather, there is just the one dignitary harm. And this is the sense in which the dignitary harm of being exposed to risk of harm is *absorbed* into the outcome harm. Yet, things are not quite so simple. Plausibly, we want to say: first, that the disrespect harm is greater, the greater the risk of harm you were subjected to, other things being equal; and, second, the disrespect harm is the same, whether the harm materialises or not, other things being equal (as Placani writes, 'before such actualization occurs [of the risked harm materialising] dignitary interests are set back by exposure to the risk'.)⁵⁸ But this means, going back to our case in which Threatener subjects Victim₁ and Victim₂ to two different risks of being electrocuted and both risked harms materialise, that Victim₁ is harmed more than Victim₂. This is because they are both harmed to the same degree by the electrocution, excluding any setbacks to dignity, and then they have different disrespect harms additionally (to the extent that Victim₁ was subjected to a greater risk of harm, and this is more disrespectful). But now we are back to the Magnitude of Harm Problem, where Victim₁ is harmed more than Victim₂.

This section has shown that, if risk is harmful, then two agents who are subjected to different levels of risk of the same outcome

⁵⁸ Placani, 'When the Risk of Harm Harms', *supra* note 3 at 88.

harm are, if both risked harms materialise, harmed to different extents all things considered. This is counterintuitive. This was the Magnitude of Harm Problem. We then considered the Absorption Solution to this problem – when risked harms materialise, the risk harm is absorbed into the outcome harm. Yet, this was even more implausible for it implies, when all else is equal, outcome harms are different depending on the level of risk one was exposed to.

VI. MOVING FORWARD SAFELY

This paper has answered two questions. First, assuming that risk is of moral significance, does risk's moral significance lay in its being harmful? And, second, is risk harmful? In reply to our second question, I have argued risk is not harmful. We have just seen that the Magnitude of Harm Problem implies that any view on which risk is harmful leads to implausible results. We have also seen that the Desire Explanation fails, because the desire that we not be at risk of harm is a derivative desire, and the setting back of derivative desires does not *itself* constitute a harm. And we have seen that the Autonomy Explanation leaves us with implausible asymmetries: exposing autonomous individuals to risk is, in some circumstances, harmful, while exposing individuals with undeveloped, compromised, or damaged rational capacities to risk is not harmful; and it says that Locker's exposing Chooser to risk is harmful in *Two Doors* but that Target's exposing Target to risk in *Roulette* is not harmful.

In reply to our first question (whether risk's moral significance lays in its being harmful), since risk is not harmful and yet risk is of moral significance, risk's moral significance cannot lie in its being harmful. However, I did note two more conservative conclusions along the way. Even if people sometimes desire not to be subject to risk of harm for non-derivative reasons, and the frustration of these desires are themselves harmful, risk's moral significance cannot lay in its being harmful on this view, since risk will not be of moral significance to those who desire not to be at risk for merely derivative reasons. And, even if a defender of the Autonomy Explanation is willing to accept all the objections and restrictions raised against it, this shows that risk's moral significance cannot lay in its being harmful, because then we could not explain why exposing those

without autonomy to risk is itself harmful, or have a satisfying explanation of *Roulette*.

So where does the moral significance of risk lie, if not in its being harmful? Strictly speaking, this goes beyond the scope of this paper. One thing we might believe is that what we ought to do and/or what rights we have is determined by what the evidence says is best.⁵⁹ How views like these would deal with risk should be relatively straightforward. For example, in *Roulette*, because what is presumably the best available evidence says there is a one in six chance of Shooter's killing Target, and nothing to gain by that, it is evidence-relative wrong for her to play Russian roulette with Target.

Let me sketch a more novel, different way we could go. It explains the moral significance of risk without moving to evidence-relativity. It takes seriously the idea that Shooter *could* have harmed Target. One way of thinking about this is that there is a relevantly close counterfactual world in which Target is harmed – namely, the world in which there is a bullet in the chamber of the gun when Shooter pulls the trigger. Because of this, Shooter's subjecting Target to risk of harm makes Target's wellbeing less safe than it would have been, had she not played roulette with Target. And, we might think that agents have rights not merely against *actually* being harmed, but that they could not easily have been harmed – that their well-being is *safely* protected across relevantly close counterfactual worlds.⁶⁰

The idea behind modal safety is nicely explained by Timothy Williamson:

Imagine a ball at the bottom of a hole, and another balanced on the tip of a cone. Both are in equilibrium, but the equilibrium is stable in the former case, unstable in the latter. A slight breath of wind would blow the second ball off; the first ball is harder to shift. The second ball is in danger of falling; the first ball is safe. Although neither ball did in fact fall, the second could easily have fallen; the first could not.⁶¹

⁵⁹ Zimmerman, *Ignorance and Moral Obligation*, *supra* note 7; Jonathan Quong, *The Morality of Defensive Force* (Oxford: Oxford University Press, 2020).

⁶⁰ Recall earlier that I objected to Lazar's view that, when one subjects another to risk, this 'contravenes [an] important interest [...] in *being secure*'. Lazar, 'Risky Killing', *supra* note 3 at 7. Does my focus on *safety* look a lot like Lazar's focus on security? In some sense, it does: we both think it matters that people could not easily be harmed. Where we disagree is whether safety is partly constitutive of wellbeing. A rough reason for preferring my view is that the modally demanding good to which Lazar and I are appealing is 'Agents' wellbeing is securely/safely protected'. This locution implies that *this* good cannot itself be partly constitutive of wellbeing, otherwise the definition would be self-referential. More substantively, having security as partly constitutive of wellbeing makes the view vulnerable to the Magnitude of Harm Objection, raised in section V. For more reasons in favour of my account, see my 'Robust Rights and Harmless Wronging', in *Oxford Studies in Normative Ethics*, ed. Mark Timmons (Oxford University Press, Forthcoming).

⁶¹ Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000), 123.

There is a danger an event will occur if that event does occur in some sufficiently similar case. And, much like how the ball is not safely balanced on the top of the cone, Target's wellbeing is not safely protected.

Suppose that one endorses the Interest Theory of Rights – the view on which individuals' claim-rights, and the correlative duties that those rights entail, are grounded in right-holders' wellbeing (her interests).⁶² One way that we might respond to the idea that individuals' well-being should be safely protected is to endorse the following condition, in-keeping with the Interest Theory:

The Safety Condition. For X to have a right against Y that $Y \Phi$, Y 's not Φ -ing must cause X to be worse off than she would have been in at least one close world, and the difference in X 's wellbeing must be of sufficient weight to place Y under a duty to Φ .

In *Roulette*, the Safety Condition is satisfied: Target is worse off in the close possible world in which Shooter plays roulette with her, and shoots her, than she is when Shooter does not play Russian roulette with her; and, the extent to which she is worse off is of sufficient weight to place Shooter under a duty not to play roulette with her. While Shooter does not end up harming Target given the way that things turn out, she does violate Target's rights.

There is obviously a lot more to be said about this account, which I develop elsewhere.⁶³ The purpose of this paper has been to argue that risk is not itself harmful and, hence, that the moral significance of risk does not lie in risk's itself being harmful.

ACKNOWLEDGMENTS

My thanks to audiences at the University of St Andrews, University of Glasgow, and Society for Applied Philosophy's Annual Conference in 2018. For comments and discussion, my thanks to Antony Duff, Romy Eskens, Helen Frowe, Thomas McKenna, Colin McLean, Sandra Marshall, Matthew Oliver, Tom Parr, Justin Snedegar, Saranga Sudarshan, and Victor Tadros. A special thanks to Stephanie Bowen, Rowan Cruft, Theron Pummer, and to Katherine Hawley, who gave me comments on this when she really didn't have to, who will be sorely missed. This work was supported by an AHRC Scottish Graduate School for the Arts and Humanities

⁶¹ Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000), 123.

⁶² For the canonical statement of the view, see Raz, *The Morality of Freedom*, *supra* note 30 at 166.

⁶³ Bowen, 'Robust Rights and Harmless Wronging', *supra* note 60.

doctoral training partnership studentship (Grant No. 1804818), a Society for Applied Philosophy Doctoral Scholarship, and finally the Knut and Alice Wallenberg Foundation (Grant No. 1521101).

FUNDING

Open access funding provided by Stockholm University.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Stockholm University, Stockholm, Sweden
E-mail: joseph.bowen@philosophy.su.se

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.